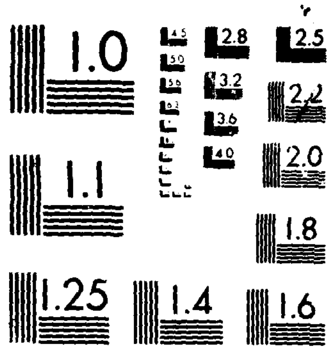
AD-A186 388 CONTROL CHARTS WHEN THE OBSERVATIONS ARE CORRELATED(U) 1/1
PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS
P R KRISHNAIAH ET AL MAY 87 TR-87-09 AFOSR-TR-87-1109
UNCLASSIFIED F49600-85-C-0008 F/G 12/3 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A186 388

DTIC FILE COPY

(2)

AFOSR-TR- 87-1109

aiah and B. Q. Miao	8. CONTRACT F-49620
NIZATION NAME AND ADDRESS ultivariate Analysis Thackeray Hall f Pittsburgh, Pittsburgh, PA 15260	10. PROGRAM E AREA & WO 61102F
ICE NAME AND ADDRESS fice of Scientific Research NM f the Air Force Force Base, DC 20332 Bldg 410	12. REPORT DA May 198 13. NUMBER OF 12
CY NAME & ADDRESS(If different from Controlling Office)	15. SECURITY C

DTIC ELECTE
OCT 0 1 1987
S D
D

Center for Multivariate Analysis

University of Pittsburgh

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

87 9 24 201

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 87-1109	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Control charts when the observations are correlated		5. TYPE OF REPORT & PERIOD COVERED Journal ~~Technical~~ - May 1987
		6. PERFORMING ORG. REPORT NUMBER 87-09
7. AUTHOR(s) P. R. Krishnaiah and B. Q. Miao		8. CONTRACT OR GRANT NUMBER(s) F-49620-85-C-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis, Fifth Floor, Thackeray Hall, University of Pittsburgh, Pittsburgh, PA 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304 A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research NM, Department of the Air Force, Bolling Air Force Base, DC 20332 Bldg 410		12. REPORT DATE May 1987
		13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Same as 11		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Autoregressive model, control ellipse, time series, $\bar{X}$-chart.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Traditionally, control charts are based on independently normal samples, but in practice it so happens that the samples are dependent. In this review, dependent samples are considered as ARMA time series. Also multidimensional time series samples are discussed.

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(2)

CONTROL CHARTS WHEN THE OBSERVATIONS ARE CORRELATED*

P.R. Krishnaiah and B.Q. Miao

Center for Multivariate Analysis
University of Pittsburgh

May 1987

Technical Report No. 87-09

Center for Multivariate Analysis
Fifth Floor Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

* Research sponsored by the Air Force Office of Scientific Research (AFOSC) under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notion hereon.

CONTROL CHARTS WHEN THE OBSERVATIONS ARE CORRELATED*

ABSTRACT

Traditionally, control charts are based on independently normal samples, but in practice it so happens that the samples are dependent. In this review, dependent samples are considered as ARMA time series. Also, multidimensional time series samples are discussed.

AMS Subject classification: Primary 62N10.

Key words and phrases: Autoregressive model, control ellipse, time series, $\overline{X}$-chart.

* Research sponsored by the Air Force Office of Scientific Research (AFOSC) under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprins for governmental purposes notwithstanding any copyright notion hereon.

1. INTRODUCTION

Traditionally, quality control charts have been designed with respect to statistical criteria only, and the control methodology is based on the independence and normality of serial samples. At first the production process is assumed to be characterized by a single in-control state. For example, if the process has one measurable quality characteristic, then the in-control state will correspond to the mean of this quality characteristic when no assignable cause is present.

Now we consider the model:

$$X_t = \mu + \xi_t \tag{1.1}$$

where μ is a constant, ξ_t is an error and X_t is t-th observation. It is of interest to select the sample size, statistical characteristic of ξ_t and control limits so that the power of the test to detect a particular shift in the quality characteristic and the type I error probability are equal to specified values. Usually, ξ_t, $t = 1, 2, \ldots$ is considered to be independently normally distributed with zero mean and common variance σ^2, where σ^2 is known or unknown. In this case, consideration of statistical criteria and practical experience have led to general guidelines for the design of control charts resulting in widespread use of samples of size 5, three-sigma control limits, and a sampling frequency of one hour for the $\overline{X}$-chart (see Duncan [3]). In the sequel, we set

$$\overline{X}_i = \frac{1}{n}\sum_{j=1}^{n} X_{in+j}, \qquad S_i = \left(\frac{1}{n}\sum_{j=1}^{n}(X_{in+j} - \overline{X}_i)^2\right)^{1/2},$$

$$\overline{\overline{X}} = \frac{1}{k}\sum_{i=1}^{k}\overline{X}_i, \qquad \overline{S} = \frac{1}{k}\sum_{i=1}^{k} S_i,$$

[illegible] the number of samples. Hereafter we use $\overline{X}$ [illegible] the general 3σ quantity control limits on [illegible]

$$[illegible] \overline{X} \pm A\sigma, \quad A = \frac{3}{\sqrt{n}}, \tag{1.2}$$

$$[illegible] \overline{X} \pm A_1\overline{S}, \quad A_1 = \frac{A}{C_2}\overline{S}, \tag{1.3}$$

[illegible] the standard deviation is known or estimated. [illegible] limits on the standard deviation are

$$[illegible] (B_1, B_2)\sigma, \tag{1.4}$$

$$[illegible] (B_3, B_4)\overline{S}, \tag{1.5}$$

[illegible] known or estimated, where

$$[illegible] \tag{1.6}$$

$$[illegible] \left([illegible] - \frac{1}{n} - C_2^2 \right)^{1/2}, \tag{1.7}$$

$$[illegible] \quad B_2 = C_2 + 3C_3, \tag{1.8}$$

$$[illegible] \quad B_4 = 1 + \frac{3C_3}{C_2}. \tag{1.9}$$

[illegible] A, A_1, $B_1 - B_4$ are tabulated in the literature (see Grant [4]). The process will be considered under control if the estimate of the mean and the estimate of the standard deviation of the process remain within prescribed control limits above.

In practice, a number of data sets in economics, business, engineering and the natural science often are present in the form of time series. In other words, the observations are dependent, i.e., ξ_t's of model (1.1) are not white noise; for example, $\{\xi_t, t = 0, \pm 1, \ldots\}$ is an autoregressive moving average (ARMA) with order (p,q). So the problem is how to determine 3σ control limits. Stamboulis [7] studied AR(1) with parameter α. Vasilopoulos [8] extended Stambouli's results to ARMA(p,q) model. Vasilopoulos and Stamboulis [9] together investigated the case AR(2) = ARMA(2,0). It is different from classical control factors. How different it is depends on the stochastic properties of the process. Since the method is similar, we only discuss AR(2).

2. CONTROL CHART ON THE SECOND ORDER AUTOREGRESSIVE MODEL

In model (1.1), assume ξ_t is an AR(2) model, that is

$$\xi_t = \alpha_1 \xi_{t-1} + \alpha_2 \xi_{t-2} + \varepsilon_t \tag{2.1}$$

where $\{\varepsilon_t\}$ is a white noise series with $E\varepsilon_t = 0$ and $V(\varepsilon_t) = \sigma_\varepsilon^2$, α_1 and α_2 are constants. For stationarity of AR(2), it is necessary that the roots of the characteristic equation of the AR(2) process

$$\phi(B) = 1 - \alpha_1 B - \alpha_2 B^2 = 0 \tag{2.2}$$

must lie outside the unit circle, which is equivalent to the alpha coefficients being in the triangular region:

$$\alpha_2 + \alpha_1 < 1, \qquad \alpha_2 - \alpha_1 < 1, \qquad -1 < \alpha_2 < 1$$

(see Box and Jenkins [2]). The variance of the AR(2) process is given by

$$\sigma^2 = \left(\frac{1 - \alpha_2}{1 + \alpha_2}\right) \frac{\sigma_\varepsilon^2}{[(1 - \alpha_2)^2 - \alpha_1^2]} . \tag{2.3}$$

Suppose γ_k, $k = 0, \pm 1, \ldots$, are the autocovariance functions of the AR(2) process, then σ^2 and the variance $\sigma_{\bar{X}}^2$ of the sample mean $\bar{X}$ are given in terms of γ_k by

$$\sigma^2 = \gamma_0, \tag{2.4}$$

$$\sigma_{\bar{X}}^2 = \frac{1}{n} [\gamma_0 + 2 \sum_{t=1}^{n-1} (1 - \frac{1}{n}) \gamma_t]. \tag{2.5}$$

In order to evaluate the control limits for $\bar{X}$, we need to evaluate $\sigma_{\bar{X}}^2$. This can be accomplished by (2.3)-(2.5) if α_1, α_2 and σ_ε^2 are known. Therefore, if the process variance σ^2 is known, the control limits of the model described by (1.1) and (2.1) is modified to

$$\bar{x} \pm A(\alpha_1,\alpha_2,n)\sigma \tag{2.6}$$

where

$$A(\alpha_1,\alpha_2,n) = \lambda^{1/2}(\alpha_1,\alpha_2,n)\frac{3}{\sqrt{n}}, \tag{2.7}$$

$$\lambda(\alpha_1,\alpha_2,n) = 1+2\sum_{t=1}^{n-1}(1-\frac{t}{n})b_t, \tag{2.8}$$

$$b_t = \gamma_t/\gamma_0. \tag{2.9}$$

The expression of $\lambda(\alpha_1,\alpha_2,n)$ is based on the expression of $\sigma^2_{\bar{x}}$.

In order to construct the $\bar{x}$-chart when the standard deviation is unknown, we must evaluate the auxiliary parameter $C_2(\alpha_1,\alpha_2,n)$ first, which is also needed to construct control limits of the standard deviation. Since S^2/ES^2 is distributed as χ^2_n, we get

$$E(S) = C_2(\alpha_1,\alpha_2,n)\sigma = \sqrt{\frac{2}{n}}\,\frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})}\left(1-\frac{2}{n-1}\sum_{t=1}^{n-1}(1-\frac{t}{n})b_t\right)^{1/2}\sigma. \tag{2.10}$$

Hence,

$$C_2(\alpha_1,\alpha_2,n) = C_2\cdot\left(1-\frac{2}{n-1}\sum_{t=1}^{n-1}(1-\frac{t}{n})b_t\right)^{1/2}. \tag{2.11}$$

To obtain an approximate expression for E(S), we use this expression of E(S):

$$E(S) = E\sqrt{S^2} = \sqrt{ES^2}\left(1-\frac{\mathrm{Var}\ S^2}{8(ES^2)^2}\right) - \frac{\sqrt{ES^2}}{8(n-1)}\left\{2n-3-\left(\frac{2\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})}\right)^2\right\}. \tag{2.12}$$

In general, the last term of (2.12) is smaller than the first term. For example, let n = 5. Then the last term of (2.11) is $0.002144\cdot\sqrt{ES^2}$, so we can omit this term and get the approximate formula:

$$E(S) \approx \sqrt{ES^2}\left(1 - \frac{\mathrm{Var}\ S^2}{8(ES^2)^2}\right). \tag{2.13}$$

The expected value and variance of S^2 can be obtained by

$$ES^2 = \gamma_0\left(1 - \frac{\lambda(\alpha_1,\alpha_2,n)}{n}\right) \tag{2.14}$$

and

$$\mathrm{Var}(S^2) = \frac{1}{n^2}\sum_{t=1}^{n}\sum_{\tau=1}^{n}\gamma_{t-\tau}^2 + 2\left[\frac{1}{n^2}\sum_{t=1}^{n}\sum_{\tau=1}^{n}\gamma_{t-\tau}\right]^2 - \frac{4}{n^3}\sum_{t=1}^{n}\sum_{\tau=1}^{n}\sum_{\gamma=1}^{n}\gamma_{t-\tau}\gamma_{t-\gamma}. \tag{2.15}$$

But, from the expression of $\mathrm{Var}(S^2)$, the complexity of (2.13) is not better than (2.10).

By replacing $A_1(\alpha_1,\alpha_2,n)$, $C_3(\alpha_1,\alpha_2,n)$, $B_i(\alpha_1,\alpha_2,n)$ for A_1, C_3 and B_i, $i = 1, 2, 3, 4$ of (1.3) and (1.6)-(1.8), respectively, the modification of control chart limits in an AR(2) model is obtained.

The substantial ranges in the values of $\lambda(\alpha_1,\alpha_2,n)$ and $C_2(\alpha_1,\alpha_2,n)$ greatly affect the control factors. Vasilopoulos and Stamboulis [9] gave an example to illustrate this result.

3. CONTROL CHARTS IN MULTIVARIATE CASE

Now we consider multivariate case. The model in this case is

$$\underset{\sim}{X}_t = \underset{\sim}{\mu} + \underset{\sim}{\xi}_t \tag{3.1}$$

where $\underset{\sim}{\mu}$ is a m×1 constant vector, $\{\xi_t,\ t = 1, 2, \ldots\}$ is a m-dimensional stationary process with zero mean vector. Set

$$\overline{\underset{\sim}{X}} = \frac{1}{n}\sum_{i=1}^{n} \underset{\sim}{x}_i, \qquad S = \frac{1}{n-1}\sum_{i=1}^{n}(\underset{\sim}{x}_i - \overline{\underset{\sim}{x}})(\underset{\sim}{x}_i - \overline{x})',$$

where the prime means transpose of a matrix (or vector). Let $\overline{\overline{\underset{\sim}{x}}}$ be the global mean over several subgroups of size n, and $\overline{S}$ be the pooled sample covariance. It is well-known that if $\{\xi_t\}$, t = 1, 2, ... is a series consisting of white noise with distribution $N_m(\underset{\sim}{0},\Lambda)$, the $\overline{\underset{\sim}{X}}$-control chart has been studied by many authors (for example, Ghare and Torgrersen, Jackson and et al). From the facts that $(\overline{\underset{\sim}{x}} - \underset{\sim}{\mu})'(\frac{1}{n}\Lambda)^{-1}(\overline{\underset{\sim}{x}} - \underset{\sim}{\mu})$ and $(\overline{\underset{\sim}{x}} - \underset{\sim}{\mu})'(\frac{1}{n}S)^{-1}(\overline{\underset{\sim}{x}} - \underset{\sim}{\mu})$ are distributed as χ_m^2 and Hotelling T^2-statistics respectively, we can construct the quality control region on $\overline{x}$ based on Λ is known or unknown. When Λ is known, the control region is

$$D = \{\overline{\underset{\sim}{x}}:\ n(\overline{\underset{\sim}{x}} - \overline{\overline{\underset{\sim}{x}}})'\Lambda^{-1}(\overline{\underset{\sim}{x}} - \overline{\overline{\underset{\sim}{x}}}) \le \chi_m^2(\alpha)\}. \tag{3.2}$$

This is an elliptical region. When Λ is unknown, the control region is

$$D = \{\overline{\underset{\sim}{x}}:\ \frac{n-m+1}{m}(\overline{\underset{\sim}{x}} - \overline{\overline{\underset{\sim}{x}}})'\overline{S}^{-1}(\overline{\underset{\sim}{x}} - \overline{\overline{\underset{\sim}{x}}}) \le F_{m,n-m+1}(\alpha)\}. \tag{3.3}$$

But in practice, $\{\xi_t\}$ is not generally a white noise series. When $\{\xi_t\}$ is serially dependent and described by p-dimensional ARMA(p,q) model, the control region of mean vector will be modified. For simplicity, we will discuss p-dimensional AR(P) model. Let

$$\underset{\sim}{X}_t = \underset{\sim}{\mu} + \underset{\sim}{\xi}_t, \tag{3.4}$$

$$\underset{\sim}{\xi}_t = B_1\underset{\sim}{\xi}_{t-1} + \dots + B_p\xi_{t-p} + \underset{\sim}{\varepsilon}_t. \tag{3.5}$$

where B_i, $i = 1, \dots, p$, are m×m matrices and $\underset{\sim}{\varepsilon}_t$ a white noise series with distribution $N_m(\underset{\sim}{0},\Sigma)$, $\Sigma > 0$. Furthermore,we can also generalize (3.5) to the following form:

$$\xi_t = B_1\xi_{t-1} + \dots + B_p\xi_{t-p} + A\underset{\sim}{\varepsilon}_t, \tag{3.6}$$

where A: m×r, $\varepsilon_t \sim N_r(\underset{\sim}{0},\Sigma)$ such that $A\Sigma A' > 0$. The model described by (3.5) and (3.6) is often met. For example, in the production of synthetic fiber the tensile strength x_1 and diameter x_2 may be equally important quality characteristics. Their flucuations mainly result from moisture, then, in proper productive process, $(x_1,x_2)'$ may be described by (3.5) and (3.6). Here we only discuss the model (3.5) because the method treating model (3.6) is the same as (3.5).

It is well-known that the necessary condition that the AR(P) model (3.5) is stationary is all the roots of determinant of $(\lambda^p I - \lambda^{p-1}B_1 - \dots - \lambda B_{p-1} - B_p)$ lie within the unit circle. Set

$$\Lambda_k \triangleq E(\underset{\sim}{x}_t - \underset{\sim}{\mu})(\underset{\sim}{x}_{t+k} - \underset{\sim}{\mu})',$$

then, there also exist "multivariate Yule Walker" equation:

$$\Lambda_0 = B_1\Lambda_1 + \dots + B_p\Lambda_p + \Sigma, \tag{3.7}$$

$$\Lambda_k' = B_1\Lambda_{k-1}' + B_2\Lambda_{k-2}' + \dots + B_p\Lambda_{k-p}', \quad k \geq 1, \tag{3.8}$$

$$\Lambda - 1 = \Lambda_k'. \tag{3.9}$$

Hence, if B_i, $i = 1, \ldots, p$, and Σ are identified from data, then all Λ_k, $k = 0, 1, \ldots$, can be calculated from (3.7)-(3.9). Furthermore, the covariance matrix of $\bar{\underset{\sim}{x}}$ can be obtained:

$$\Lambda_{\bar{x}} = \frac{1}{n}\left(\Lambda_0 + \sum_{t=1}^{n-1}(1 - \frac{1}{n})(\Lambda_t + \Lambda_t')\right) , \tag{3.10}$$

where Λ_0 is the covariance matrix of $\underset{\sim}{x}_t$.

Since $(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu})'\Lambda_{\bar{x}}^{-1}(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu})$ is a χ^2 distribution with degrees of freedom m, we can get the control region of mean vector within an elliptical region:

$$D = \{\bar{x}: (\bar{x} - \bar{\bar{x}})'\Lambda_{\bar{x}}^{-1}(\bar{x} - \bar{\bar{x}}) \leq \chi_m^2(\alpha)\}. \tag{3.11}$$

Notice that if $\underset{\sim}{\xi}_t = \underset{\sim}{\varepsilon}_t$ in (3.5), i.e., our process is classical, then the control region described by (3.11) is the same as the one described by (3.2). When $p \neq 0$, these elliptical regions described by (3.2) and (3.11) are different from the lengths and directions of their major axes.

REFERENCES

[1] ALT, F.B. and SMITH, N.D. (1987). Multivariate process control. In *Handbook of Statistics*, Vol. 7: *Quality Control and Reliability*, North-Holland.

[2] BOX, G.E.P. and JENKINS, G.M. (1970). *Time Series Analysis: Forecasting and Control.* Holden-Day, San Francisco.

[3] DUNCAN, A.J. (1974). *Quality Control and Industrial Statistics*, 4th ed., Richard D. Irwin, Homewood, Illinois.

[4] GRANT, G.L. (1964). *Statistical Quality Control*, 3rd ed., McGraw-Hill, New York.

[5] JACKSON, J.E. (1959). Quality control methods for several related variables. *Technometrics*, **1**, 359-377.

[6] KRISHNAIAH, P.R. and MIAO, B.Q. (1987). Unpublished manuscript.

[7] STAMBOULIS, A.P. (1971). First order autoregressive model applied to quality control. New York University memorandum.

[8] VASILOPOULOS, A.V. (1974). Second order autoregressive model applied to quality control, Ph.D. dissertation, New York University.

[9] VASILOPOULOS, A.V. and STAMBOULIS, A.P. (1978). Modification of control chart limits in the presence of data correlations. *Journal of Quality Technology*, **10**, 20-30.

END
DATE
FILMED
JAN
1988